

COMP 790-173 : Module 3 Assignment - Report

Annie Rajkumar, Ian Thomas, William Su

Precondition

1. Learning

The Oxford Dictionary defines **energy** as “power derived from the utilization of physical or chemical resources, especially to provide light and heat or to work machines”. Energy can be derived from various sources grouped under two main types- renewable and non-renewable energy. In simple words, when the resource used to produce energy has the ability to eventually run out, it's labeled non-renewable, while resources that are continuously replenished by nature are called renewable. Greenhouse gas emission, a major by-product of the usage of non-renewable energy sources, is the leading cause of climate change, therefore requiring immediate action in the form of policy change.

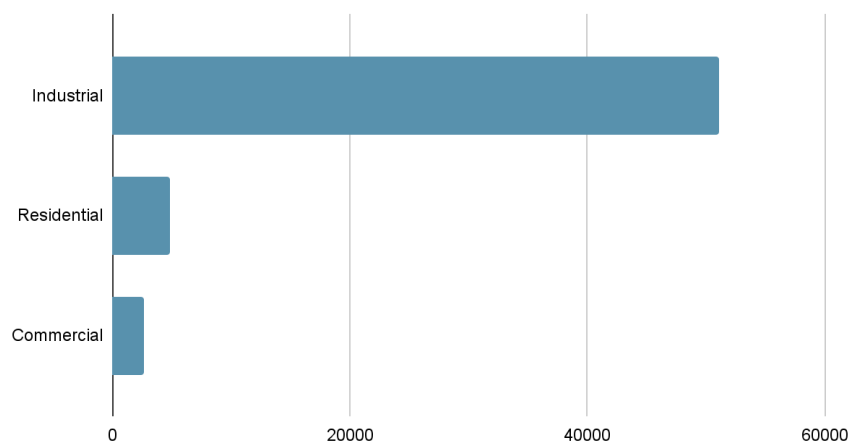
The US Energy Information Administration, in the State Energy Data System (SEDS), has collected comprehensive state energy statistics over the past 60 years. Through this data, our team set out to find which states are making better progress in switching to renewable sources. We also wanted to find and focus on the sector that is consuming the most non-renewable energy, and thereby contributing negatively to climate change, and use the insights from this data to more restrictive policies surrounding energy usage in these sectors.

2. Winnowing

The dataset we used consisted of various features related to both renewable (geothermal, hydropower, solar, wood, wind) and non-renewable (coal, distillate fuel oil, kerosene, petroleum, natural gas) energy sources, their usage in different sectors (industrial, transportation, residential, commercial, electric power, refinery), and their expenditure, price, and consumption across the 50 states in the US between 1960 and 2019.

To find the focus of one of our questions, we decided to do a baseline analysis of the data to find the sector with the highest non-renewable energy expenditure. This led us to focus specifically on petroleum expenditure in the industrial sector.

Maximum Yearly Petroleum Expenditure by Sector



To find the answer to our question on which states are making better progress towards switching to renewable energy sources, we decided on using the data on energy consumption in the industrial sector across the years in each state, and how this compares with the price of a unit of that energy source over the years. By comparing the consumption of the energy with the price per unit, we hoped to see if there was a financial reason for the usage of non-renewable energy sources versus renewable energy sources.

To proceed with our data analysis and the subsequent creation of our dashboard, we first had to manually assign each energy source as either renewable or non-renewable. There were also multiple dimensions we needed to explore in the form of comparison between different states, and time series analysis for energy usage over time. The final data consisted of 3061 rows and 24 columns. The columns were: State, Year, as well as industrial price and consumption of the 11 energy sources.

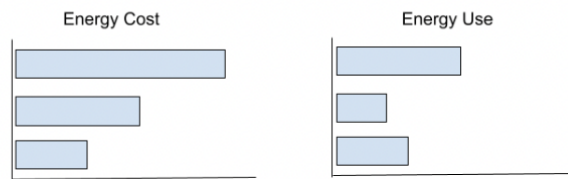
Core

3. *Discover*

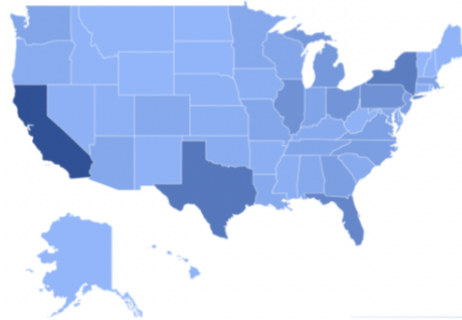
The tasks we decided to focus on are:

1. How has energy usage across states improved over time - in terms of moving away from non-renewable energy and switching to greener alternatives?
 - Who is using the most non-renewable energy and the most renewable energy?
 - What is causing this amount of usage or lack of usage of renewable energy?
 - When did the usage of renewable energy start gaining steam?
 - Why are some states using more renewable energy sources as compared to others?
 - Where should policy be implemented to start making a change?
2. How has the price of a unit of energy influenced the usage of various energy sources across different states?
 - Who is being influenced to use more renewable or non-renewable energy by its pricing?
 - What is influencing the usage and pricing of energy across different states?
 - Was there a time when the price of an energy source corresponded directly to the increase or decrease of its consumption?
 - Where did the price of energy play a direct role in its consumption?
 - Why can these trends inform better policy decisions?

4. *Design*



Charts for Energy usage



U.S. Heatmap, by state



Input for changing time

5. Implement

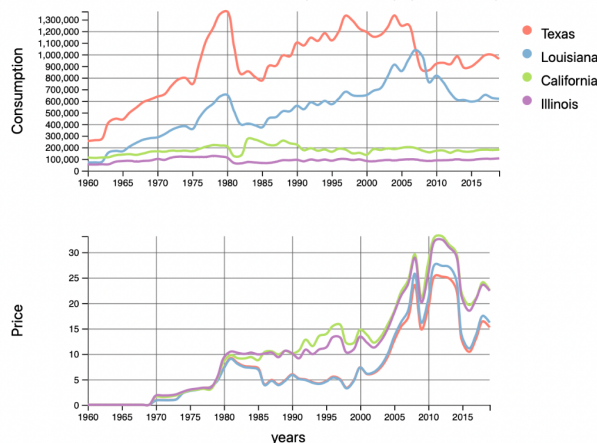
Energy Analytics Dashboard

By Annie, William, Ian

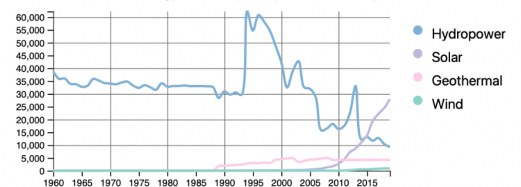
Energy Source

Other Petroleum Products

Other Petroleum Products Consumption & Price (Top Four States)



Renewable Energy Consumption (Total, All States)

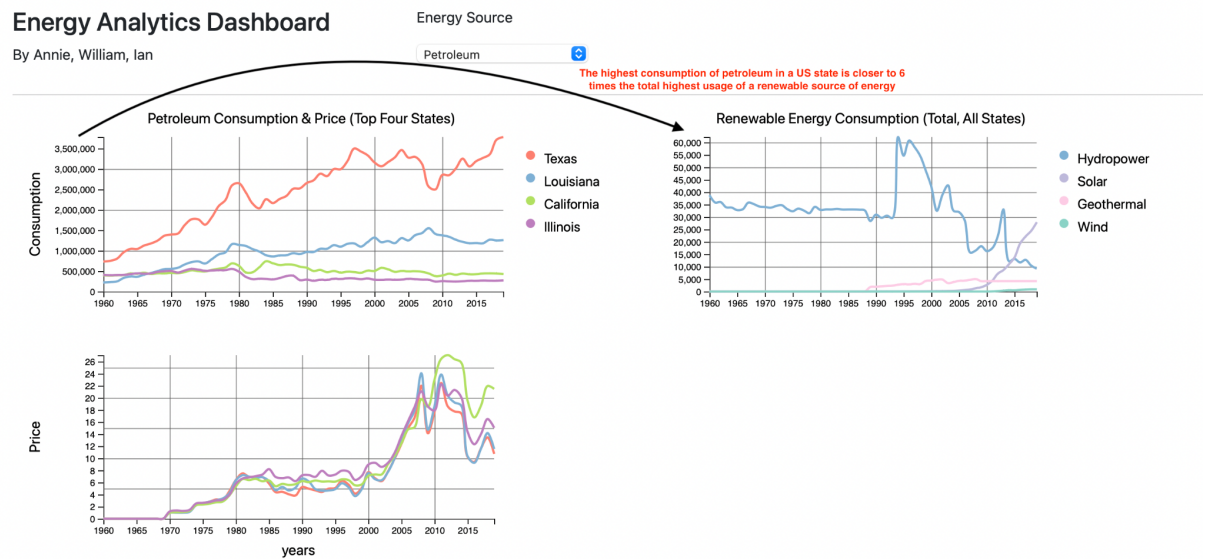


The geographic visualization was ultimately dropped as the only information it would provide to the user is the physical location of states (which if this utility is being used to inform policy change, it is assumed the user is aware of state locations).

Instead, all data is time-series data for energy source usage over time, with the user being able to filter by energy source. The dashboard only shows the top four states as showing more is cognitively intense on the user, and almost all other states, in the case of most energy sources, are so close together that the difference is negligible.

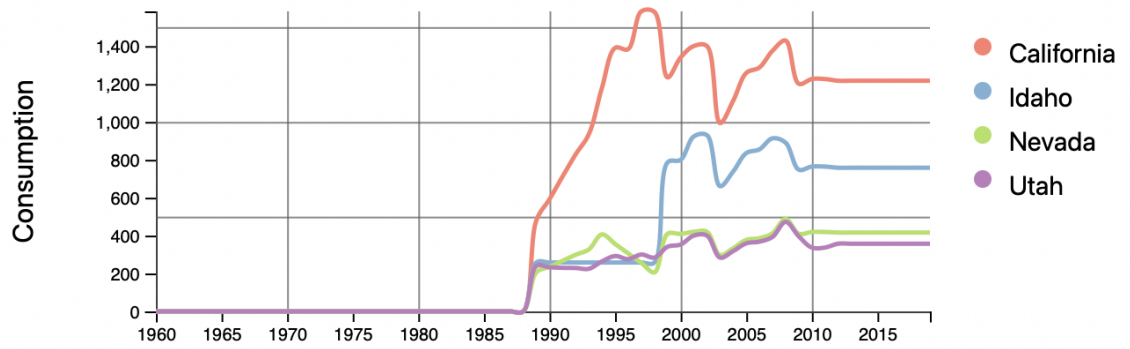
6. Deploy

Observation 1: The usage of non-renewable energy even when broken down by state is much higher than the total usage of renewable energy across the US.

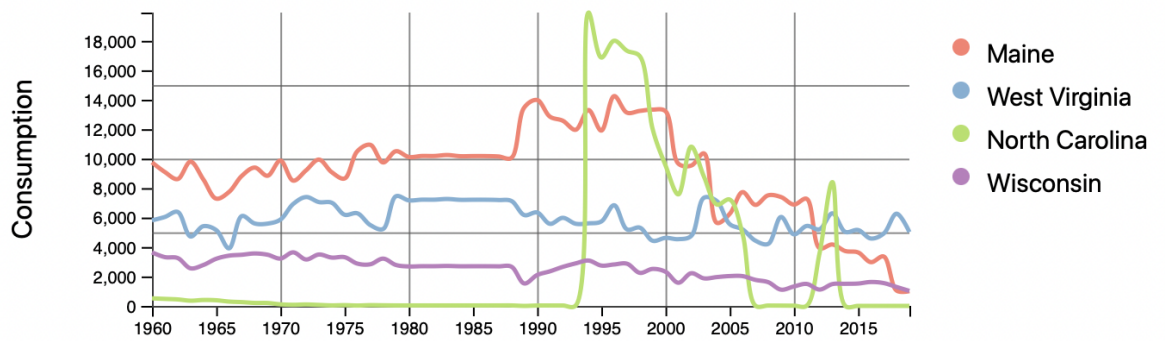


Observation 2: No two states (except California) show high consumption of more than one type of renewable energy source, which could speak to the availability of said resource in this state. For example, California ranks first in both geothermal and solar energy usage probably due to its year-round sunny weather. On looking at the consumption axis, it can be seen that none of their total renewable energy usage comes close to the usage of non-renewable energy.

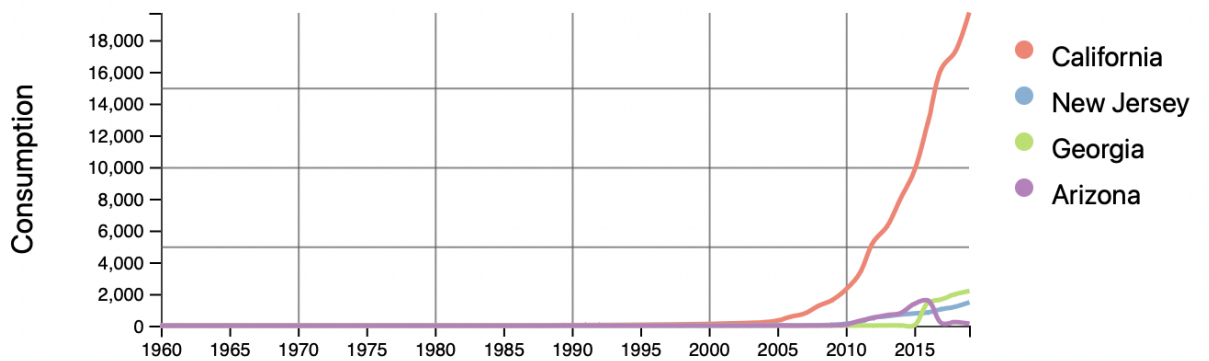
Geothermal Consumption & Price (Top Four States)

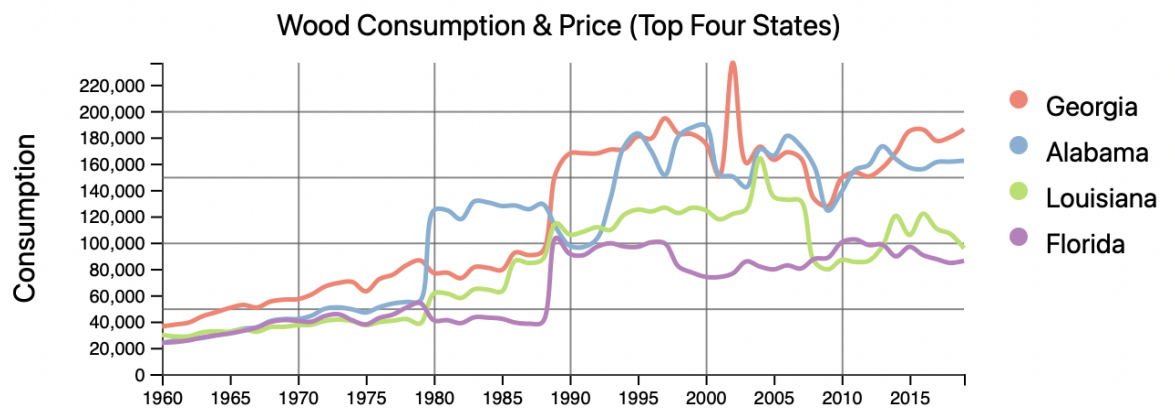
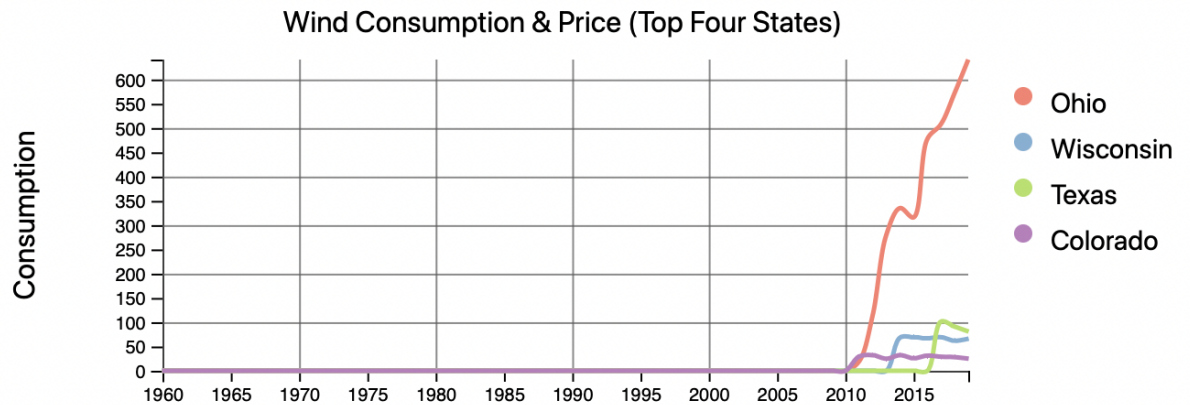


Hydropower Consumption & Price (Top Four States)

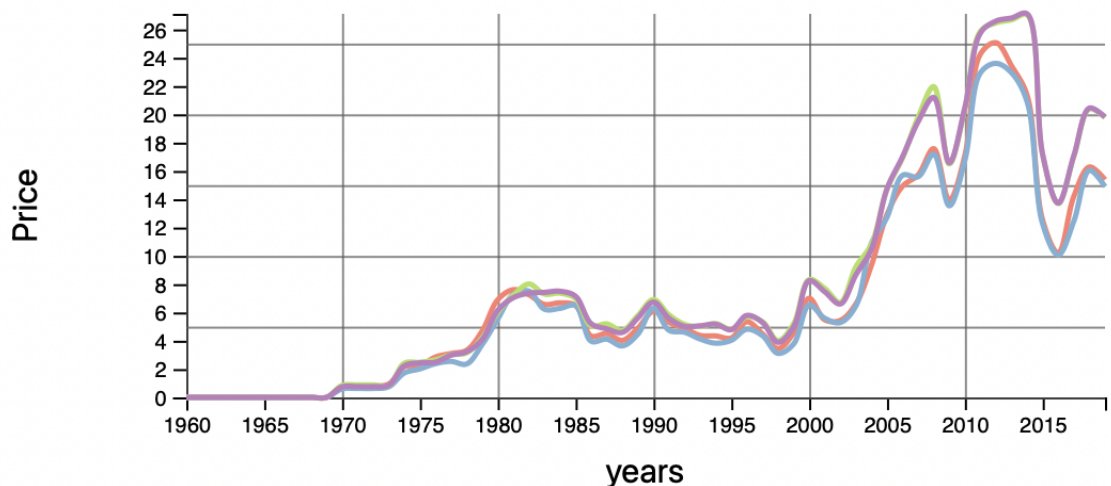


Solar Consumption & Price (Top Four States)

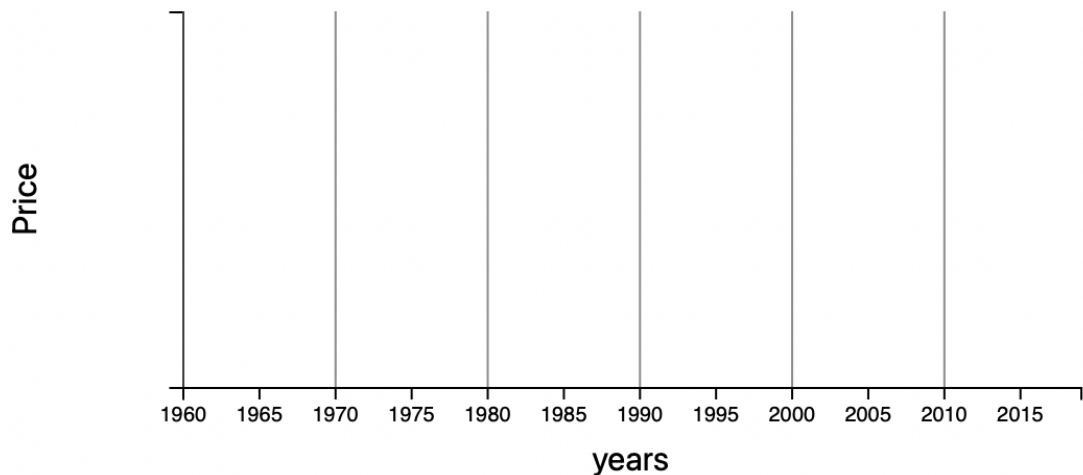




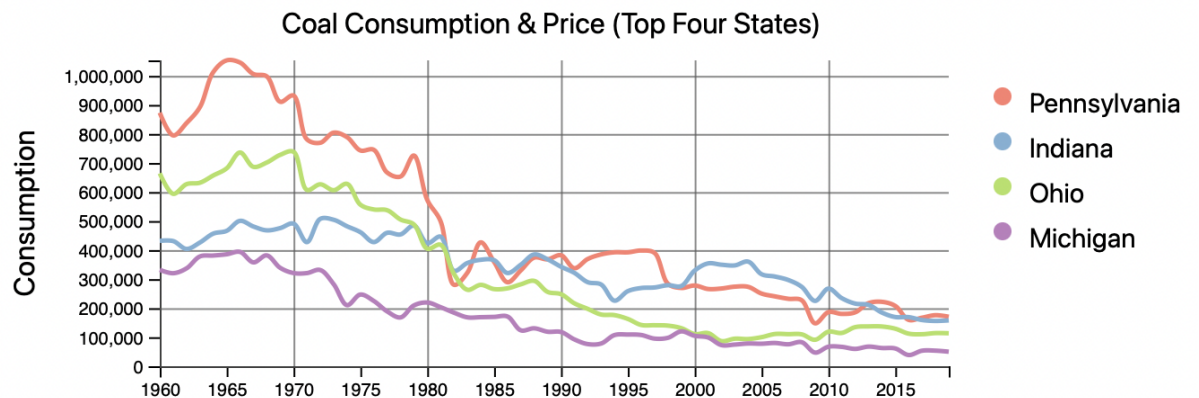
Observation 3: Prices of non-renewable energy sources have changed quite equally across all states, therefore leading to our conclusion that there is no significant relationship between the usage of non-renewable energy and its price per unit.



Observation 4: The absence of the prices of renewable sources of energy over time. This led to not being able to estimate whether its price over time is the reason for its low usage.



Observation 5: Coal usage has reduced over time.



7. Iterate

We could perhaps add more data relevant to why states have used more energy than they have. For example, data on population and GDP could be a good indicator as to why a state has been better or worse at switching energy sources.

Analysis

8. Reflect, Pt 1

Our solution provides some key insights into which states have been the largest consumers of certain energy sources. The system could provide policymakers with insights on which states have performed better or worse at switching to renewable energy sources. Currently, the Energy Policy Act of 2005 requires that in every fiscal year, the federal government consumes at least 7.5% of its electricity from renewable sources, and the American Renewable Energy Act of 2021 has established a goal of requiring electricity suppliers to achieve at least 70% renewable electricity generation by 2030, therefore requiring recommendations to be made across all sectors and states on how they can transition to cleaner energy sources over the next couple of years.

9. *Reflect, Pt 2*

This solution tells us that this data set is missing some key components in providing insights for evaluating how well states have performed in switching to renewable energy sources. For example, the lack of price for renewable energy sources makes it harder to assess how well some states have performed as there is no way to evaluate the financial and economical impact. This solution also tells us that sometimes a certain data point can skew results causing the visualization to be unusable for our purposes, and that the data point causing the skew might have to be removed from the visualization and analyzed separately to allow for a better understanding of all the other data. For example, while creating our chart on the total consumption of renewable energy sources, we found that due to the high usage of wood, all the other renewable sources were barely visible. This caused us to exclude wood and instead focus the chart on the other renewable sources.